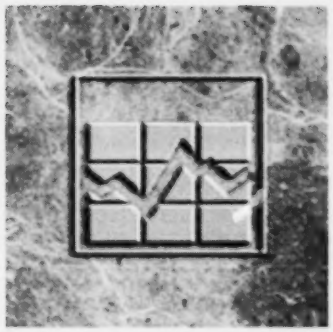
N° 75F0002MIF au catalogue — N° 007
ISSN : 1707-2867
ISBN : 0-662-72279-5

Document de recherche

Série de documents de recherche - Revenu

# Données sur le logement dans l'Enquête sur la dynamique du travail et du revenu (EDTR)



par Heather Lathe

Division de la statistique du revenu
Immeuble Jean-Talon, Ottawa, K1A 0T6

Téléphone: 1 613 951-7355

Statistique Canada Statistics Canada

Canada

## Comment obtenir d'autres renseignements

Toute demande de renseignements au sujet du présent produit ou au sujet de statistiques ou de services connexes doit être adressée à : Division de la statistique du revenu, Statistique Canada, Ottawa, Ontario, K1A 0T6 (téléphone : (613) 951-7355; (888) 297-7355: revenu@statcan.ca).

Pour obtenir des renseignements sur l'ensemble des données de Statistique Canada qui sont disponibles, veuillez composer l'un des numéros sans frais suivants. Vous pouvez également communiquer avec nous par courriel ou visiter notre site Web.

| | |
|---|---|
| Service national de renseignements | 1 800 263-1136 |
| Service national d'appareils de télécommunications pour les malentendants | 1 800 363-7629 |
| Renseignements concernant le Programme des services de dépôt | 1 800 700-1033 |
| Télécopieur pour le Programme des services de dépôt | 1 800 889-9734 |
| Renseignements par courriel | infostats@statcan.ca |
| Site Web | www.statcan.ca |

## Renseignements pour accéder au produit

Le produit n° 75F0002MIF au catalogue est disponible gratuitement. Pour obtenir un exemplaire, il suffit de visiter notre site Web à www.statcan.ca et de choisir la rubrique Nos produits et services.

## Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois, et ce, dans la langue officielle de leur choix. À cet égard, notre organisme s'est doté de normes de service à la clientèle qui doivent être observées par les employés lorsqu'ils offrent des services à la clientèle. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1 800 263-1136. Les normes de services sont aussi publiées sur www.statcan.ca sous À propos de Statistique Canada > Offrir des services aux Canadiens.

Statistique Canada
Division de la statistique du revenu

Série de documents de recherche - Revenu

# Données sur le logement dans l'Enquête sur la dynamique du travail et du revenu (EDTR)

Publication autorisée par le ministre responsable de Statistique Canada

© Ministre de l'Industrie, 2006

Tous droits réservés. Le contenu de la présente publication électronique peut être reproduit en tout ou en partie, et par quelque moyen que ce soit, sans autre permission de Statistique Canada, sous réserve que la reproduction soit effectuée uniquement à des fi ns d'étude privée de recherche, de critique, de compte rendu ou en vue d'en préparer un résumé destiné aux journaux et/ou à des fi ns non commerciales. Statistique Canada doit être cité comme suit : Source (ou « Adapte de », s'il y a lieu) : Statistique Canada, année de publication, nom du produit, numéro au catalogue, volume et numéro, période de référence et page(s). Autrement, il est interdit de reproduire le contenu de la présente publication, ou de l'emmagasiner dans un système d'extraction, ou de le transmettre sous quelque forme ou par quelque moyen que ce soit, reproduction électronique, mécanique, photographique, pour quelque fi n que ce soit, sans l'autorisation écrite préalable des Services d'octroi de licences, Division des services à la clientèle, Statistique Canada, Ottawa, Ontario, Canada K1A 0T6.

Juillet 2006

N° 75F0002MIF au catalogue, vol. 7

ISSN : 1707-2867
ISBN : 0-662-72279-5

Périodicité : hors série

Ottawa

This publication is available in English upon request (Catalogue no. 75F0002MIE).

**Note de reconnaissance**

*Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population, les entreprises, les administrations canadiennes et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques précises et actuelles.*

## Résumé

Le présent document résume les données sur les caractéristiques du logement et les dépenses liées au logement qui sont disponibles dans l'Enquête sur la dynamique du travail et du revenu (EDTR), une attention particulière étant portée aux méthodes d'imputation utilisées. De 1994 à 2001, l'enquête ne portait que sur quelques caractéristiques, surtout le mode d'occupation et le type de logement. En 2002, grâce à la commandite de la Société canadienne d'hypothèques et de logement (SCHL), plusieurs autres caractéristiques ainsi que des dépenses détaillées liées au logement ont été ajoutées à l'enquête. Diverses méthodes d'imputation ont aussi été adoptées à ce moment-là, pour remplacer les valeurs manquantes attribuables à la non-réponse à l'enquête, et pour fournir les coûts des services publics qui contribuent au coût total du logement. Ces méthodes tirent parti du plan de sondage longitudinal de l'EDTR, et elles utilisent également des données provenant d'autres sources comme l'Enquête sur la population active et le Recensement. En juillet 2006, d'autres améliorations aux méthodes d'imputation ont été adoptées pour 2004 et appliquées à des années antérieures dans le cadre d'une révision historique. Le présent rapport documente également cette révision.

## Remerciements

L'auteure désire remercier de leurs contributions et commentaires Chris Duddek et Susan Storey à la Division des méthodes d'enquêtes auprès des ménages, ainsi que Steve Grunau, Willa Rea, Preston Poon et Anne-Marie Vinette à la Division de la statistique du revenu.

## Aperçu

1. Introduction ..... 7
2. Participation de la Société canadienne d'hypothèques et de logement ..... 9
3. Résumé des données disponibles ..... 9
4. Remplacement des valeurs manquantes : Imputation longitudinale ..... 12
5. Remplacement des valeurs manquantes : Imputation transversale ou du « donneur ». ..... 14
6. Révision des données sur le logement des années 1999 à 2004 ..... 15
7. Taux d'imputation et résultats de la révision ..... 16
8. Imputation des coûts des services publics ..... 21
9. Où trouver plus d'information ..... 23
Références ..... 25

# 1. Introduction

L'Enquête sur la dynamique du travail et du revenu (EDTR) est une enquête-ménage menée tous les ans par Statistique Canada au moyen d'interviews téléphoniques. Elle est conçue pour mesurer le bien-être socio-économique des Canadiens, y compris les changements au fil du temps et les facteurs qui influent sur ces changements.

Les caractéristiques du logement et les dépenses liées au logement des Canadiens sont devenues, assez récemment, l'une des principales composantes de l'EDTR. Les données sur les revenus des ménages (ou des familles), sur les caractéristiques démographiques, sur la participation au marché du travail et sur diverses autres composantes font partie de l'enquête depuis sa mise en œuvre en 1993. La plupart des variables sur le logement ont été mises au point à compter de l'année de référence 2002 (elles ont été recueillies en janvier 2003) à la demande de la Société canadienne d'hypothèques et de logement (SCHL).[1] La SCHL avait besoin de données annuelles sur le revenu et le logement, qui seraient recueillies au moyen d'une enquête menée auprès d'un échantillon relativement nombreux de répondants, pour compléter les données du Recensement de la population (tous les cinq ans) et de l'Enquête sur les dépenses des ménages. La capacité de procéder à une analyse longitudinale au moyen des données de l'EDTR présentait également de l'intérêt pour la SCHL. Au moment de la rédaction du présent rapport, cette commandite annuelle du contenu sur le logement dans l'EDTR se poursuit chaque année.

Le présent rapport résume les données sur les caractéristiques du logement et les dépenses liées au logement qui sont disponibles dans l'EDTR, une attention particulière étant portée aux méthodes d'imputation utilisées pour ces données. L'imputation joue un rôle important dans la plupart des enquêtes-ménages, car elle constitue une façon de remplacer ce qui serait autrement des valeurs manquantes dans les données de sortie. Les

1. La Société canadienne d'hypothèques et de logement (SCHL) est l'organisme national responsable de l'habitation au Canada. La SCHL travaille à améliorer les options financières en ce qui a trait à l'habitation au Canada; elle aide les Canadiens à faible revenu à accéder à la propriété; elle améliore les normes en matière de construction d'habitation et donne aux responsables des politiques l'information et l'analyse dont ils ont besoin pour assurer la prospérité du marché de l'habitation au Canada.

valeurs manquantes découlent de non-réponses ou d'erreurs dans les données recueillies auprès des répondants. Même si on tient compte d'une non-réponse complète au questionnaire au moyen de la redistribution des poids de l'enquête, une non-réponse à certaines questions peut être réglée pendant le traitement de l'enquête grâce à l'imputation (l'attribution) de valeurs valides appropriées au lieu de réponses « ne sait pas ».

Dans le cas des valeurs manquantes, l'imputation est utilisée dans l'EDTR pour toutes les variables relatives au revenu et au logement. En outre, l'EDTR utilise l'imputation pour couvrir les dépenses au titre de services publics, comme l'électricité, le chauffage et l'eau, qui sont difficiles à recueillir à un degré élevé d'exactitude lorsqu'on utilise des interviews téléphoniques. Ce type d'imputation est appelé imputation générale, car il est utilisé pour obtenir les dépenses liées aux services publics pour 100 % des répondants dans l'échantillon.

La suite de ce document présente d'abord un bref aperçu du rôle de la SCHL dans l'enquête. Ensuite, on décrit brièvement le contenu, avec des explications sur la façon de consulter le site Web de Statistique Canada de l'information supplémentaire figurant dans d'autres publications. Les sections 4 et 5 décrivent dans des termes non techniques les divers types d'imputation utilisés pour remplacer les valeurs manquantes dans les variables sur le logement de l'EDTR. Une explication plus technique se trouve dans un document antérieur, intitulé *Imputation générale des variables du logement (sans les services publics) dans l'Enquête sur la dynamique du travail et du revenu* (quoique certains des tableaux ne soient plus à jour). Les sections 6 et 7 décrivent comment l'imputation a été réappliquée pour les années 2002 à 2004 peu après la publication des données de 2004, dans le but d'offrir une amélioration. Ces sections décrivent également comment ces mêmes méthodes ont été appliquées rétroactivement à 1999 pour les variables qui existaient avant 2002. La section 8 décrit brièvement comment l'imputation est utilisée pour imputer les coûts des services publics pour tous les ménages dans le but d'inclure ces coûts dans les estimés du coût total du logement, étant donné que les coûts exacts des services publics sont difficiles à recueillir par téléphone.

## 2. Participation de la Société canadienne d'hypothèques et de logement

La Société canadienne d'hypothèques et de logement (SCHL) évalue les conditions du logement et développe des estimés des besoins en logement en utilisant des données extraites toutes les cinq ans du Recensement de la population. Pour analyser les conditions du logement entre les recensements, la SCHL utilise des données d'enquêtes, par exemple des informations des données de l'Enquête sur les dépenses des ménages (EDM) et avant 1997, l'Enquête sur l'équipement ménager (EEM). Cela lui permet d'avoir des tendances annuelles et un plus grand choix d'autres caractéristiques pour le recoupement des données sur le logement des ménages canadiens.

En 2001, la SCHL a commencé à commanditer des questions supplémentaires à la fois dans l'EDTR et dans l'EDM, à compter de l'année de référence 2002. Les questions supplémentaires ont été tout particulièrement importantes pour l'EDTR. Jusqu'alors, l'EDTR n'avait porté que sur quelques caractéristiques, surtout le mode d'occupation et le type de logement. De plus, aucune imputation des caractéristiques du logement n'a été faite. Statistique Canada a ajouté plus de 20 questions sur le logement à l'interview de l'EDTR. Certaines questions s'adressent aux propriétaires-occupants et aux locataires, certaines uniquement aux propriétaires-occupants, et d'autres uniquement aux locataires. Seulement une personne répond à ces questions pour le compte de tous les membres du ménage.

## 3. Résumé des données disponibles

Cette section renferme une liste des caractéristiques du logement disponibles à des fins d'analyse. Chaque élément ou « variable » est accompagné du nom abrégé de la variable utilisé pour la programmation, d'un bref label pour mieux identifier ce qu'elle signifie, et de la question (ou des questions) qui servent à la recueillir dans le questionnaire. En outre, des descriptions plus longues des variables et des descriptions des catégories figurent sur Internet dans le dictionnaire de données de l'EDTR. Le libellé précis des questions et l'ordre dans lequel elles sont posées pendant l'interview figurent dans le

document de recherche qui renferme le module entrée-sortie du questionnaire. Voir la section 9, « Où trouver plus d'information ».

Comme il est expliqué ci-dessus, la majorité des questions sur le logement dans l'EDTR ont été introduites en 2002, grâce à la commandite de la SCHL.

La première année de référence de l'EDTR, 1993, ne contenait aucune variable sur le logement. Pour les années de référence 1994 à 1998, deux questions ont été ajoutées, soit le mode d'occupation (DWLTEN25) et le type de logement à quatre catégories (DWLTYP25). À compter de 1999, deux nouvelles questions ont été ajoutées, en partie en raison de la normalisation entre les enquêtes du questionnaire assisté par ordinateur. Les nouvelles questions portaient sur le nombre de chambres dans le logement, si le propriétaire était membre du ménage et le cas échéant, si le propriétaire avait une hypothèque.

Également à compter de l'année de référence 1999, les catégories de réponses pour la question sur le type de logement ont été augmentées de quatre (maison individuelle, maison jumelée, maison en rangée et appartement) à sept, comme dans la variable DWLDET25. Pour de nombreuses fins analytiques, il suffit d'avoir seulement les quatre grandes catégories de types de logement et, par conséquent, DWLTYP25 est une version abrégée de DWLDET25 depuis 1999.

| Nom de la variable | Label | Numéro de la question |
|---|---|---|
| **Disponible depuis 1999:** | | |
| DWLDET25 | Type détaillé du logement habité (9 catégories) | DWT_Q01; CON-Q6C avant 1999 |
| DWLTYP25 | Type de logement (4 catégories) | Dérivée; CON-Q6D avant 1999 |
| DWTENR25 | Logement possédé ou loué | TN_Q01 |
| ROOMS25 | Nombre de chambres à coucher | BD_Q01 |
| MORTG25 | Hypothèque sur le logement (oui/non) | MTGE_Q01 |
| **Disponible depuis 2002:** | | |
| REPA25 | Réparations nécessaires au logement | OR_Q005 |
| HEAT25* | Principal combustible utilisé (aussi, HEATG25) | OR_Q010 |
| OPFM25 | Ménage exploite une ferme sur cette propriété | OR_Q015 |
| OPBU25** | Ménage exploite une entreprise sur cette propriété | OR_Q020 |
| PRTXI25** | Impôts fonciers inclut dans les paiements hypothécaires | OW_Q005 |
| PRTXM25 | Valeur mensuelle des impôts fonciers | OW_Q045 |
| MORTGN25 | Plus d'un prêt hypothécaire sur le logement | OW_Q010 |
| MORTGP25** | Fréquence des paiements hypothécaires | OW_Q015 |
| MORTGM25 | Paiements hypothécaires mensuels, excl. impôts fonciers | OW_Q025 – Q040 |
| COND25 | Logements en condominium | OW_Q050 |
| CONDM25 | Montant des frais de copropriété mensuels | OW_Q055 |
| RENTM25 | Loyer mensuel (excl. les services publics) | RN_Q005 |
| RNPK25 | Stationnement inclus dans le loyer | RN_Q010 |
| RNHT25 | Chauffage inclus dans le loyer | RN_Q010 |
| RNWA25 | Eau incluse dans le loyer | RN_Q010 |
| RNEC25 | Électricité incluse dans le loyer | RN_Q010 |
| RNTV25* | Câble inclus dans le loyer | RN_Q010 |
| RNFG25* | Réfrigérateur inclus dans le loyer | RN_Q010 |
| RNST25* | Cuisinière incluse dans le loyer | RN_Q010 |
| RNWD25* | Laveuse/sécheuse incluses dans le loyer | RN_Q010 |
| RNFU25 | Meubles inclus dans le loyer | RN_Q010 |
| RNNO25 | Aucun élément inclus dans le loyer | RN_Q010 |
| FURN25* | Type d'ameublement inclus dans le loyer | RN_Q010 |
| RNRE25 | Loyer mensuel subventionné (oui/non et raison) | RN_Q020 |
| RNBS25 | Loyer payé calculé en fonction du revenu (oui/non) | RN_Q025 |
| UHTY25, UWAY25, UECY25 | Coût du combustible, d'eau et d'électricité, respectivement | Toutes valeurs imputées |
| SHELY25 | Coûts annuels d'habitation | Dérivée par addition |
| RQNU25, SUIT25 | Nombre de chambres à coucher requises pour le ménage, et si ce nombre est suffisant. | Dérivée*** |

* Disponible jusqu'en 2005 seulement

** Valeurs manquantes due à une non-réponse sont laissées telles quelles dans les données de sorties pour les années 2002 – 2005, mais seront remplacées par des valeurs imputées et ce, débutant en 2006.

*** De la part de la Société canadienne d'hypothèques et de logement, «Études spéciales sur les données du recensement de 1996 : conditions de logement au Canada ». Le point en recherche : Série socio-économique, Numéro 55-1, p.1, 4 : Un logement est adéquate si le nombre de chambres à coucher dans le logement est égal ou supérieur au nombre de chambres à coucher requises pour le ménage selon la Norme nationale d'occupation.

À compter de l'année de référence 2006, deux nouvelles questions ont été ajoutées. Les deux sont importantes pour l'imputation des dépenses liées aux services publics, qui sont incluses dans le coût total du logement. Dans une nouvelle question, on demande aux propriétaires si les paiements pour l'électricité, le combustible de chauffage ou l'eau sont

inclus dans d'autres paiements comme les frais de copropriété. En supposant que les frais de copropriété déclarés comprennent déjà ces paiements, nous pouvons imputer 0 $ pour ces services publics. Une nouvelle question s'adresse aux ménages locataires qui déclarent 0 $ pour les paiements mensuels de loyer, où on leur demande s'ils paient pour des services publics et, dans l'affirmative, lesquels. Dans ce cas, nous imputons pour ces ménages un montant positif pour les services publics.

Également à compter de l'année de référence 2006, la question sur le type de logement est subdivisée en environ quatre brèves questions pour recueillir l'information au téléphone d'une façon plus conviviale (l'EDTR est une enquête téléphonique, alors que l'EPA utilise un mélange d'interviews téléphoniques et d'interviews « sur place », c.-à-d. en face à face). La longue liste antérieure de catégories, qui devait être lue au répondant au téléphone n'était pas perçue favorablement[2].

## 4. Remplacement des valeurs manquantes : Imputation longitudinale

On utilise l'imputation longitudinale dans les enquêtes où la collecte des données comporte un aspect longitudinal, de sorte que les mêmes répondants sont interviewés pour plus d'une période de collecte. Dans la mesure du possible, on tient compte de l'information déjà recueillie auprès du répondant au cours d'un contact antérieur. L'EDTR suit les répondants pendant une période pouvant aller jusqu'à six ans. Ses répondants participent également à l'Enquête sur la population active (EPA) un an avant de participer à l'EDTR — ce qui donne une autre occasion d'utiliser l'imputation longitudinale.

L'utilité de l'imputation longitudinale est grandement accrue lorsque les caractéristiques à imputer sont censées demeurer les mêmes d'une période à l'autre, selon certaines autres

2. Il se peut qu'elle ait également donné lieu à une proportion trop élevée de réponses à la toute première catégorie, « maison individuelle », comparativement au Recensement et à l'EDM qui n'utilisent pas un mode de collecte téléphonique. Lorsque les données de 2006 de l'EDTR deviendront disponibles, on pourra déterminer si la question modifiée aligne davantage les résultats de l'EDTR concernant le type de logement avec les résultats du Recensement et de l'EDM, ou si un problème sous-jacent du biais de non-réponse en est la cause.

caractéristiques. Par exemple, nous pouvons tirer parti du fait qu'un grand nombre des caractéristiques du logement devraient être les mêmes si le répondant n'a pas déménagé. On peut utiliser le code postal pour déterminer, au moins avec une probabilité raisonnable, que le répondant n'a pas déménagé.

On obtient le code postal du répondant de son adresse, chaque fois qu'il est interviewé. S'il y a eu un contact avec le répondant pendant une année particulière, il est probable que l'on a obtenu le code postal, même si l'interview n'a pas été réalisée au complet. La plupart des répondants qui n'ont pas déjà été contactés font l'objet d'une redistribution des poids entre les ménages des répondants, de sorte que leurs poids finals sont zéros, ce qui signifie qu'ils sont exclus des estimés et qu'il n'est donc pas nécessaire de procéder à une imputation dans leur cas.

L'EPA et le questionnaire sur le loyer, qui est un supplément de l'EPA, ont été utilisés comme sources d'information pour l'imputation longitudinale dans le cas des répondants qui ont participé pour la première année à l'EDTR. Les données manquantes de l'EDTR ont été imputées à partir de l'information EPA sur le ménage, si celui-ci n'avait pas déménagé. Les quatre variables qui pouvaient être imputées de cette façon étaient le type de logement, le mode d'occupation, le nombre de chambres à coucher, et (pour les propriétaires) s'il y avait une hypothèque.

Pour utiliser le questionnaire sur le loyer, il a fallu s'assurer que la variable mode d'occupation (propriété) dans l'EDTR indiquait que le ménage était encore un locataire. Les variables dont les valeurs sont peut-être tirées de l'Enquête sur le loyer dans l'imputation longitudinale comprennent les indicateurs des commodités (services publics, stationnement, etc.) incluses dans les paiements locatifs du ménage, et le montant du loyer mensuel.

Outre les données de l'EPA ou du questionnaire sur le loyer, les données de l'EDTR de l'année précédente sont également utilisées pour imputer les données pour les années subséquentes, si le ménage ne déménage pas. Les données imputées peuvent être reportées sur plus d'un an, mais les ménages qui ne participent pas du tout à l'enquête

pendant des années successives se voient généralement attribuer un poids de zéro pour ces années.

## 5. Remplacement des valeurs manquantes : Imputation transversale ou du « donneur »

Pour résumer, l'imputation par enregistrement donneur consiste à repérer un groupe de ménages qui partagent plusieurs caractéristiques avec le ménage dont les données doivent être imputées, et à reproduire les données d'un ménage dans la zone manquante du ménage soumis à l'imputation. Les caractéristiques communes qui sont entrées sont choisies à partir d'autres variables dans l'enquête. Ainsi, dans l'EDTR on impute le revenu d'un particulier en choisissant le revenu d'une personne du même âge, du même sexe et du même niveau d'instruction vivant dans la même province, ayant le même type d'emploi (salarié, travailleur indépendant) et la même profession.

Dans le cas des variables sur le logement de l'EDTR, les variables à imputer sont traitées dans des étapes ou des modules distincts, selon qu'elles s'appliquent à tous les ménages ou uniquement à certains types de ménages, comme ceux qui louent leur logement, ceux qui sont propriétaires, les propriétaires qui ont une hypothèque (variables reliées à l'hypothèque), etc. Les ménages ne nécessitent pas tous une imputation complète des caractéristiques du logement, auquel cas ils peuvent être traités dans seulement un ou deux des modules d'imputation. Les modules sont structurés de telle façon que même si un ménage ne nécessite pas une imputation complète des caractéristiques du logement, les variables imputées dans des modules antérieurs peuvent alors être utilisées comme entrées dans d'autres.

Le premier module sert à imputer le type de logement et le mode d'occupation. Si un ménage doit être soumis à une imputation pour l'une ou l'autre de ces variables, il doit probablement faire l'objet d'une imputation complète, car ces variables figurent parmi les premières questions posées au sujet du logement dans le questionnaire de l'EDTR. Les variables d'entrée utilisées pour regrouper les ménages donneurs et les ménages à soumettre à une imputation sont la province, le groupe de taille de région urbaine (non décrit ici) et le nombre de personnes dans le ménage. Il ne s'agit là que de la première

étape de l'imputation par enregistrement donneur sur les caractéristiques du logement dans l'EDTR. On trouve plus d'information dans le document de recherche technique intitulé *Imputation générale des variables du logement (sans les services publics) dans l'Enquête sur la dynamique du travail et du revenu.*

## 6. Révision des données sur le logement des années 1999 à 2004

La publication du présent rapport a eu lieu en même temps qu'une rediffusion des données sur le logement de l'EDTR pour les années 1999 à 2004. Le reste du contenu de l'EDTR n'a pas changé par rapport à ce qui a été diffusé le 30 mars 2006.

La révision des données sur le logement visait la réimputation des coûts des services publics. En outre, on voulait appliquer l'imputation des valeurs manquantes à un plus grand nombre d'années : les techniques elles-mêmes n'ont pas changé, mais elles ont été élargies pour s'appliquer rétroactivement à 1999, et elles ont été réappliquées de façon plus complète pour 2002-2004, afin d'accroître l'exactitude et l'utilité des données recueillies. Il a été décidé, pour des raisons opérationnelles, de ne pas appliquer les méthodes d'imputation aux années antérieures.

Une imputation longitudinale et une imputation par enregistrement donneur ont maintenant été effectuées sur les données de 1999 à 2001. Plus précisément, l'imputation longitudinale a été utilisée pour les répondants qui ont commencé à participer à l'enquête en 1999 (répondants du panel 3) et a été effectué à l'aide de leurs données de l'EPA de 1998, y compris les données du questionnaire sur le loyer de l'EPA. L'imputation longitudinale à l'aide des données de l'EDTR de l'année précédente a ensuite été appliquée aux trois années (répondants du panel 2 seulement pour 1999). On a pu également procéder à une imputation longitudinale rétrospective, c.-à-d. en utilisant des données des années subséquentes, étant donné qu'on effectuait une révision historique.

Enfin, une imputation par enregistrement donneur a été appliquée aux trois années pour remplacer les valeurs qui manquaient encore[1].

Le remplacement des valeurs manquantes de 1999 à 2001 a permis d'effectuer, pour la première fois, une imputation longitudinale pour les répondants du panel 3 en 2002, et ce, pour les variables clés du type de logement, du mode d'occupation, du nombre de chambres à coucher et de l'existence d'une hypothèque. Pendant que l'on effectuait cette imputation, on a également saisi l'occasion d'améliorer l'imputation longitudinale pour les répondants du panel 4 en 2002. Plus précisément, en raison d'une erreur dans la façon dont on a consulté à l'origine leurs données dans le questionnaire sur le loyer de 2001, le questionnaire a été sous-utilisé, mais cette erreur a maintenant été corrigée. L'imputation longitudinale a ensuite été réappliquée pour 2003 et 2004. L'augmentation résultante des données de 2002 à 2004 soumises à l'imputation longitudinale a ensuite réduit la dépendance à l'égard de l'imputation par enregistrement donneur lorsqu'elle a été réappliquée pour les trois années.

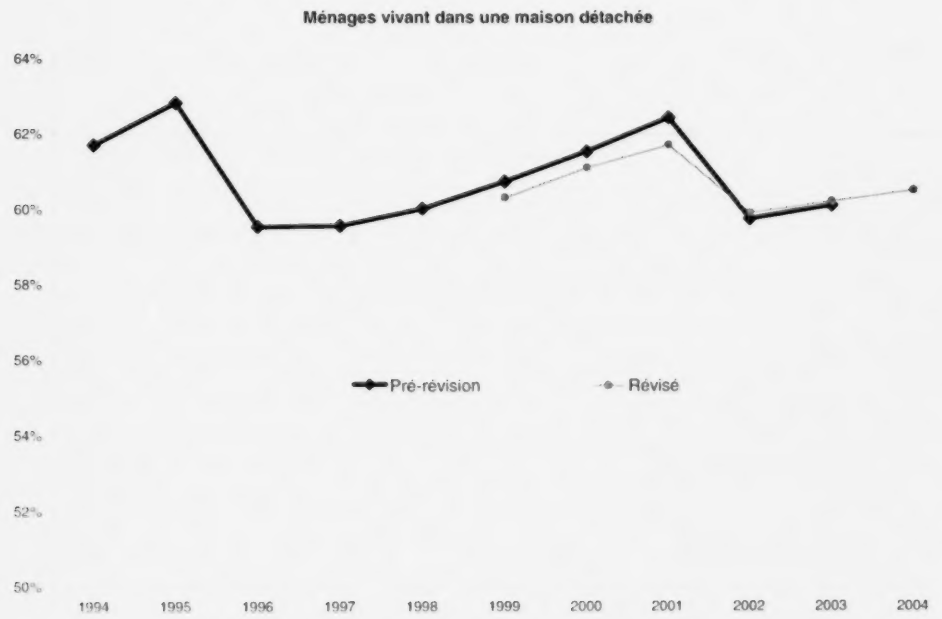
Tout en adoptant ces meilleures méthodes pour remplacer les valeurs manquantes sur le logement, deux changements additionnels ont été apportés. Premièrement, la variable indiquant si une ferme est exploitée sur la propriété (OPFM25) a été révisée pour 2003 pour qu'elle concorde avec les autres années sur le plan de la population pouvant se voir attribuer une valeur valide. Deuxièmement, on a renversé une révision de 1999 où les ménages qui ne comptaient aucune chambre à coucher (ROOMS25) avaient été modifiés à des ménages ayant une chambre à coucher.

1. À titre de note historique, l'imputation longitudinale a déjà été utilisée pour les variables sur le logement pour 1999 et 2000 depuis la diffusion de 2001 de l'EDTR. Cela a été fait pour compenser les erreurs de collecte au cours de ces années attribuables à l'informatisation du questionnaire EDTR. Si l'EDTR n'avait pas été une enquête longitudinale, les données manquantes pour le type de logement en 1999 et les quatre caractéristiques du logement en 2000 auraient touché près de 40 % des ménages. Plutôt, l'imputation longitudinale a réduit les taux de non-réponse à des niveaux beaucoup plus acceptables. On n'a pas procédé à l'imputation des données du donneur à ce moment-là.

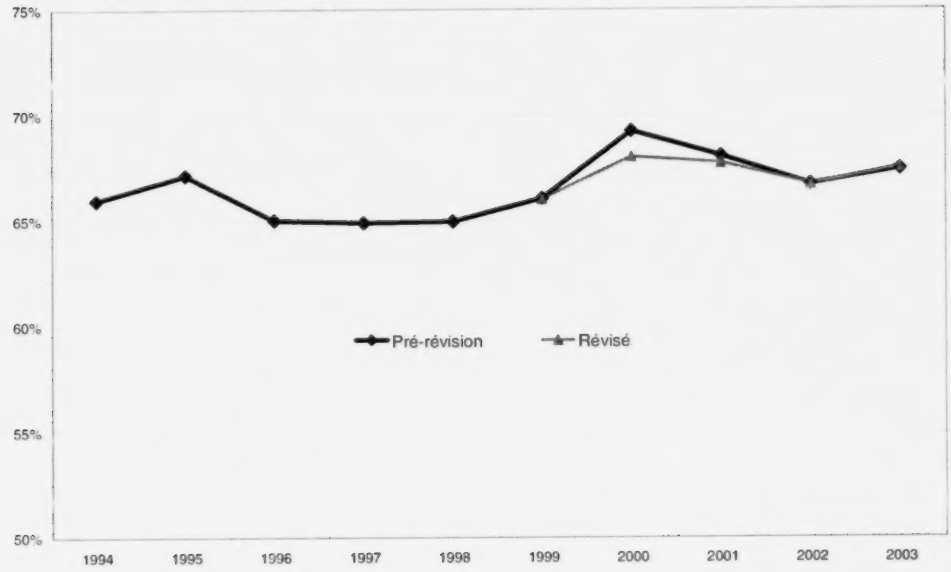
## 7. Taux d'imputation et résultats de la révision

Suite à la révision de 1999 à 2004, il n'y a plus de valeurs manquantes pour le type de logement, le mode d'occupation, le nombre de chambres à coucher et si le ménage a une hypothèque. Disposer de données complètes sans valeurs manquantes accroît l'utilité des données aux niveaux d'agrégation plus faibles, par exemple pour les petites régions géographiques, car il existe plus d'enregistrements dans l'échantillon disponibles à des fins d'estimations.

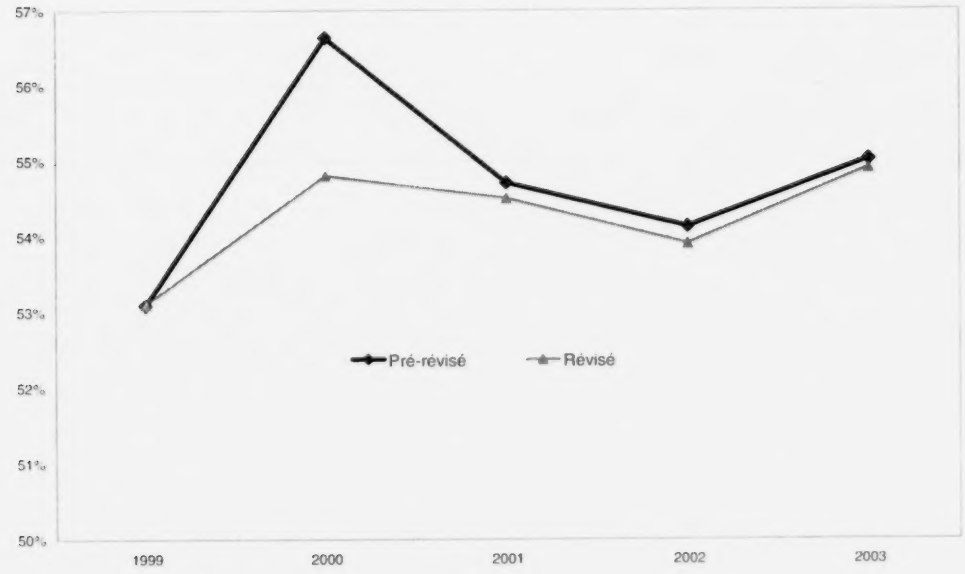
Les tableaux suivants montrent les tendances avant et après la révision dans le pourcentage de ménages vivant dans un logement individuel non attenant, de ménages propriétaires de leur logement, et de ménages qui ont une hypothèque (propriétaires seulement), respectivement. (Les années 1994 à 1998 sont incluses pour indiquer la tendance à long terme selon l'enquête, bien que les pourcentages pour 1994-1998 aient été obtenus en excluant tout simplement les enregistrements comportant des valeurs manquantes.)



**Ménages propriétaires de leur maison**




**Propriétaires ayant une hypothéque**



Si l'on se fonde sur ces trois variables, la révision a eu un impact favorable sur les données aux niveaux agrégés, particulièrement pour les quatre années de 1999 à 2002. Les nouvelles tendances pour ces trois caractéristiques ont été quelque peu lissées.

Toutefois, il y a encore des indices du biais de non-réponse, particulièrement entre 2001 et 2002, alors qu'environ la moitié des répondants de l'échantillon ont été remplacés lors de la création du panel 4 et du retrait du panel 2. L'explication la plus probable tient à l'« effet de panel » causé par l'attrition, c'est-à-dire que le taux élevé de non-réponse pendant la dernière année d'un panel est remplacé par un bien meilleur taux de réponse avec la création du nouveau panel. Même si l'imputation relative aux valeurs manquantes peut réduire l'incidence du biais de non-réponse, elle ne l'élimine pas nécessairement[4].

Le tableau 1 est une version mise à jour du tableau 8.4 présenté dans *La qualité des données dans l'Enquête sur la dynamique du travail et du revenu (EDTR) en 2004*, qui montre des taux d'imputation partiels et globaux. Toutefois, contrairement au Rapport sur la qualité des données de 2004, il exclut maintenant les ménages qui se sont vu attribuer un poids final de zéro, pour tenir compte du fait qu'une non-réponse complète à l'enquête est traitée au moyen de rajustements des poids et non de l'imputation. En outre, la variable devant indiquer si le ménage exploite une ferme sur la propriété (OPFM25) a été exclue du tableau : la révision visant à changer la variable OPFM25 à « sans objet » pour les ménages urbains, qui était auparavant comptée comme une imputation partielle, se traduisait par une surdéclaration des taux d'imputation partiels et une sous-déclaration des taux d'imputation globaux (sauf au Nouveau-Brunswick où elle s'applique pas dans le cas des ménages aussi bien urbains que ruraux).

4. On espérait que le biais de non-réponse serait encore davantage réduit par le jumelage des méthodes d'imputation longitudinales et transversales : alors que la première méthode peut être appliquée aux ménages qui n'ont pas déménagé, la deuxième peut l'être aux autres ménages. Le résultat n'était pas aussi bon qu'on l'avait prévu, mais cela peut peut-être s'expliquer par les deux facteurs suivants : premièrement, il y a relativement peu de caractéristiques dans les modules d'imputation des données du donneur; à l'heure actuelle, seules les caractéristiques au niveau du ménage sont utilisées, mais on pourrait tenir compte de certaines caractéristiques au niveau des membres dans l'avenir. Deuxièmement, l'échantillon disponible d'enregistrements de donneurs possibles a déjà été réduit par la non-réponse à l'enquête, ce qui contribue au biais de non-réponse à l'intérieur du bassin de donneurs; il serait peut-être intéressant de prendre en compte les poids finals à cette étape.

**Tableau 1 Ménages nécessitant une imputation des variables de logement, 2004**

| Province | Imputation totale | Imputation partielle | Aucune imputation |
|---|---|---|---|
| | pourcentage | | |
| Terre-Neuve | 5.2 | 27.5 | 67.3 |
| Île-du-Prince-Edward | 4.5 | 27.2 | 68.2 |
| Nouvelle-Écosse | 5.0 | 23.0 | 72.1 |
| Nouveau-Brunswick | 5.6 | 20.3 | 74.2 |
| Québec | 6.1 | 15.4 | 78.6 |
| Ontario | 7.0 | 24.1 | 68.9 |
| Manitoba | 5.5 | 26.6 | 68.0 |
| Saskatchewan | 5.0 | 27.3 | 67.7 |
| Alberta | 6.6 | 26.7 | 66.8 |
| Colombie-britannique | 7.9 | 24.7 | 67.4 |
| **Canada** | **6.2** | **23.0** | **70.8** |

Le tableau 2 montre, au moyen des exemples, à quel point l'imputation des valeurs manquantes était nécessaire, et comment les valeurs d'imputation ont été obtenues. Pour chaque variable, les ménages qui avaient fourni une valeur valide pendant l'interview n'avaient pas besoin d'imputation. Pour tous les autres ménages, on a eu recours soit à l'imputation longitudinale, soit à l'imputation par donneur. En 1999 et en 2002, l'imputation longitudinale était possible en utilisant l'EPA comme source de données pour les ménages des panels 3 et 4, respectivement. Les taux indiqués dans le tableau ont été tirés à partir des données pondérées, qui font exclusion des ménages qui étaient complètement non répondants à l'interview.

**Tableau 2 Distribution des ménages par la source des données, pour certaines variables et certaines années**

| | Tous les ménages | | Ménages qui louent | Ménages ayant une hypothèque |
|---|---|---|---|---|
| | Type de logement (DWLTYP25) | Logement possédé ou loué (DWTENR25) | Loyer mensuel (RENTM25) | Paiements hypothécaires mensuels (MORTGM25) |
| **1999** | | | | |
| Venant de la collecte | 97.10 | 53.72 | ... | ... |
| Longitudinal: EPA | 0.19 | 43.48 | ... | ... |
| Longitudinal: EDTR | 2.62 | 0.90 | ... | ... |
| Donneur | 0.09 | 1.90 | ... | ... |
| **2002** | | | | |
| Venant de la collecte | 93.15 | 93.43 | 87.91 | 70.83 |
| Longitudinal: EPA | 1.92 | 1.84 | ... | ... |
| Longitudinal: EDTR | 4.51 | 4.31 | 2.59 | 8.82 |
| Donneur | 0.42 | 0.42 | 9.50 | 20.35 |
| **2004** | | | | |
| Venant de la collecte | 92.15 | 92.30 | 85.45 | 71.87 |
| Longitudinal: EPA | ... | ... | ... | ... |
| Longitudinal: EDTR | 7.45 | 7.27 | 10.51 | 12.85 |
| Donneur | 0.40 | 0.43 | 4.04 | 15.28 |

... n'ayant pas lieu de figurer

## 8. Imputation des coûts des services publics

Les coûts des services publics (eau, combustible et électricité) peuvent représenter de grosses dépenses pour un grand nombre de ménages et jouer un rôle dans l'abordabilité du logement. Toutefois, il est difficile d'obtenir des répondants des renseignements exacts au cours d'une interview téléphonique. Plutôt que d'interviewer les répondants au sujet de ces dépenses, l'EDTR utilise une méthode appelée imputation moyenne pour ajouter cette information, en se servant des données du Recensement. Comme cette

méthode est utilisée pour l'échantillon complet de l'EDTR, on l'appelle également imputation générale.

Tous les cinq ans, le Recensement recueille des données sur les coûts des services publics pour 20 % de l'échantillon qui remplit le « questionnaire complet ». On obtient ainsi un ensemble très riche de données — surtout pour ce qui est de la précision de la collectivité — à partir desquelles on peut dégager les valeurs d'imputation. Outre la précision dans les collectivités, un grand nombre d'autres caractéristiques sont utilisées comme entrées pour imputer les dépenses au titre des services publics, c'est-à-dire : type de logement, mode d'occupation (propriété), s'il s'agit d'un condominium, le nombre de chambres à coucher, le nombre de personnes dans le ménage et le revenu du ménage.

Pour les ménages locataires, on peut également dégager une valeur à partir du Recensement, qui est fondée sur la variable indiquant si le chauffage, l'électricité et les autres services sont compris dans le loyer. À compter de l'année de référence 2006, ce rajustement pourra aussi être apporté dans le cas des ménages propriétaires, en raison de l'ajout d'une nouvelle question dans l'EDTR où l'on demandera si ces coûts sont inclus dans d'autres paiements (comme les frais de condominium ou les impôts fonciers); entre-temps, on impute une valeur positive aux coûts des services publics pour tous les ménages propriétaires.

Une fois que les ménages dans le Recensement sont regroupés avec des enregistrements semblables dans l'EDTR selon les caractéristiques d'entrée, trois sont choisis au hasard, et les coûts moyens des services publics sont ensuite utilisés comme la valeur pour le répondant à l'EDTR. S'il y a moins de trois ménages dans le Recensement qui peuvent être utilisés pour établir la valeur moyenne, les caractéristiques sont regroupées pour garantir qu'on n'utilise pas la valeur d'un seul ménage. Par exemple, les collectivités pourraient être regroupées de sorte que les sept les plus proches soient combinées.

Comme le Recensement n'est réalisé que tous les cinq ans, les valeurs moyennes de l'année du Recensement sont mises à jour annuellement au moyen d'indices provinciaux des prix payés par les consommateurs pour les services publics.

Avant la rediffusion des données sur le logement dans l'EDTR pour 1999-2004 et la publication du présent rapport, on a procédé à une imputation générale des coûts des services publics dans l'EDTR en utilisant comme source l'ensemble de données provenant de l'Enquête sur les dépenses des ménages (EDM) annuelle. Toutefois, en raison de la petite taille des échantillons, laquelle a donné lieu au retrait et au regroupement d'un grand nombre de caractéristiques d'entrée pour obtenir des valeurs moyennes, les valeurs imputées étaient trop neutres pour donner une répartition acceptable des coûts des services publics. Il a été décidé d'abandonner la source de données de l'EDM et de retenir les valeurs moyennes plus précises qui pouvaient être obtenues du Recensement; les indices des prix serviraient à mettre à jour ces valeurs moyennes entre les recensements. Par conséquent, les coûts des services publics pour 2002-2004 ont été révisés au moyen des données du Recensement.

## 9. Où trouver plus d'information

On peut trouver toute l'information relative à l'enquête sur le site Web de Statistique Canada (www.statcan.ca), à *Enquête sur la dynamique du travail et du revenu – Un aperçu de l'enquête*. Cliquer sur « Nos produits et services », « Publications Internet gratuites », « Finances personnelles et finances des ménages » et « Série de documents de recherche – Revenu ».

On trouve également à cet endroit l'*Enquête sur la dynamique du travail et du revenu – dictionnaire électronique de données*, qui renferme des données détaillées, par variable, organisées par thème. À la date du présent rapport, on peut consulter les variables sur le logement en allant aux thèmes « Caractéristiques personnelles », puis « Caractéristiques de la famille et du ménage », puis « Logement, habitation ».

La Série de documents de recherche – Revenu renferme chaque année les questionnaires complets de l'EDTR. Chercher le titre le plus récent qui commence par « Enquête sur la dynamique du travail et du revenu (EDTR) : Questionnaire de l'entrevue de la composante d'entrée et sortie ».

Pour plus de renseignements, communiquer avec les Services à la clientèle, Division de la statistique du revenu, 1 888 297-7355, télécopieur : 613 951-3012, courriel : revenu@statcan.ca.

## Références


Chris Li, Gaétan Garneau et Heather Lathe. 2005. *Modifications récemment apportées au volet géographique de l'Enquête sur la dynamique du travail et du revenu (EDTR)*. Série de documents de recherche – Revenu. No au catalogue : 75F0002MIF2005002. Ottawa : Statistique Canada.

Division de la statistique du revenu. 2006. *Enquête sur la dynamique du travail et du revenue : un aperçu de l'enquête.* Série de documents de recherche – Revenu. No au catalogue : 75F0011XIE2006001. Ottawa : Statistique Canada.

Duddek, Chris. 2006. "Modeling SLID Utility Costs using 2001 Census Data". Le 24 mai 2006. Document interne. Ottawa : Statistique Canada.

Duddek, Chris, et Susan Storey. 2006. *La qualité des données dans l'Enquête sur la dynamique du travail et du revenu (EDTR) en 2004*. Série de documents de recherche – Revenu. No au catalogue : 75F0002MIE2006005. Ottawa: Statistique Canada.

House, Georgina. 2005. *Imputation générale des variables du logement (sans les services publics) dans l'Enquête sur la dynamique du travail et du revenu (EDTR)*. Série de documents de recherche – Revenu. No au catalogue : 75F002MIE2005010. Ottawa : Statistique Canada.

Société canadienne d'hypothèques et de logement. 2000. «Études spéciales sur les données du recensement de 1996 : conditions de logement au Canada ». *Le point en recherche : Série socio-économique,* Numéro 55-1.

Storey, Susan. 2006. "Historic Housing Imputation Revision for the Survey of Labour and Income Dynamics Reference Year 1999-2004". Le 24 avril 2006. Document interne. Ottawa : Statistique Canada.